

After filter the poorly modeled Reads :

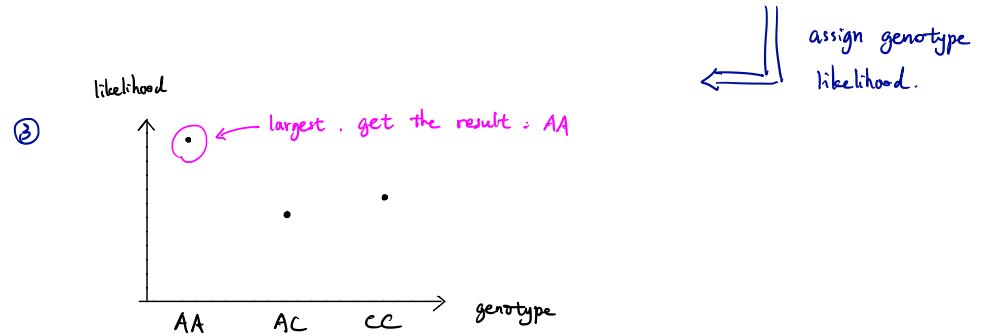
Original method :

① (Read likelihoods are exact)

	Allele : C		Allele : A	
	Hap ①	Hap ②	Hap ③	Hap ④
Read ①	1	②	3	④
Read ②	1	③	2	④
Read ③	2	③	1	④

transform to Allele-Read Table
⇒

	Allele C	Allele A
Read ①	2	4
Read ②	3	4
Read ③	3	4



Prune method :

① Read likelihoods only have bound values.

lowerbound :

	Allele : C		Allele : A	
	Hap ①	Hap ②	Hap ③	Hap ④
Read ①	1	<u>2</u>	<u>3</u>	1
Read ②	1	<u>2</u>	<u>2</u>	2
Read ③	1	<u>3</u>	<u>1</u>	1

upperbound :

	Hap ①	Hap ②	Hap ③	Hap ④
Read ①	1	<u>3</u>	<u>3</u>	2
Read ②	1	<u>2</u>	<u>2</u>	<u>4</u>
Read ③	2	<u>3</u>	1	<u>2</u>

When transform to Allele-Read Map,
still pick the best of lower and upperbound

②

lower bound :

	Allele C	Allele A
Read ①	2	3
Read ②	2	2
Read ③	3	1

upperbound :

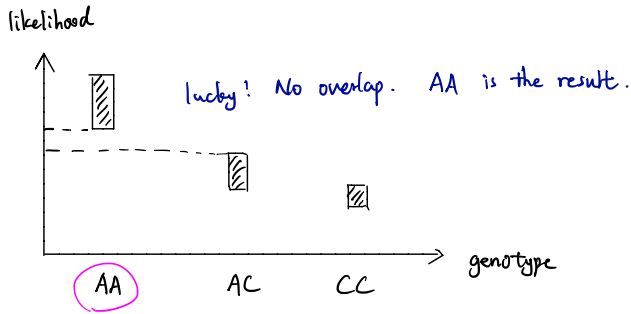
	Allele C	Allele A
Read ①	3	3
Read ②	2	4
Read ③	3	2

Note that best of lowerbound is still lowerbound
best of upperbound is still upperbound.

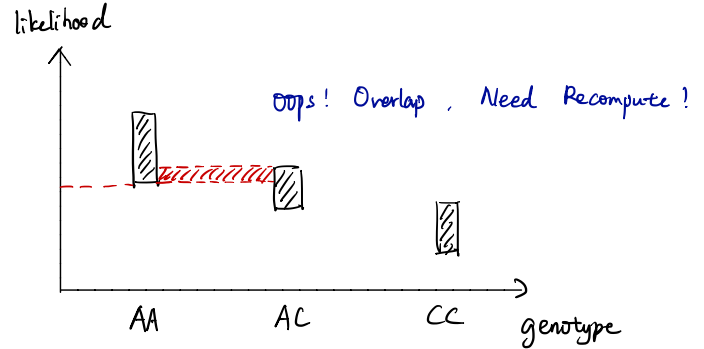
So the interval is still correct, but upperbound can be overestimated.

③ Now we assign genotype likelihoods according to last two matrices.

First case : No overlap



Second case : Overlap



↓ Recompute Step

④ Return back to the matrix in ② :

lower bound:

	Allele C	Allele A
Read ①	2	3
Read ②	2	2
Read ③	3	1

← The gap is 2!
The largest one

upperbound :

	Allele C	Allele A
Read ①	3	3
Read ②	2	4
Read ③	3	2

How to decide which Read to recompute? \Rightarrow Check the gap between lower and upperbound.
The larger the gap is, the worse this read performs. (recompute !)

So we recompute read ② against Haplotype ③ and ④ (As they have allele "A")

↓

Go back to step ③ . until it generates the exact result.